From: PETERSON Jenn L
To: Eric Blischke/R10/USEPA/US@EPA; Robert W. Gensemer
Cc: Burt Shephard/R10/USEPA/US@EPA; Chip Humphrey/R10/USEPA/US@EPA
Subject: RE: Bioassay Interpretation
Date: 06/12/2008 09:01 AM

Remember, these models are trying to predict toxicity with likely a subset of the chemicals present in the environment (or, that we are detecting at biologically relevant concentrations). This is definitely an important reason to include empirical results at the 10% / statically significant level. Even for the risk drivers identified for model development (a subset of the contamination detected in sediment), these contaminants exist in some cases in various forms of relative bioavailability as well as in different mixtures. Often, the range of conc. needed for model development is not obtained by the harbor samples. To that end, bringing in more data may improve model predictability. In the end, we need to use the predictive models as a line of evidence and understand that there is no substitute for empirical tests. This should be emphasized in PRG development, where there is a tendency to push SQGs from these models into cleanup levels.

-Jennifer

-----Original Message-----
From: Blischke.Eric@epamail.epa.gov
[mailto:Blischke.Eric@epamail.epa.gov]
Sent: Wednesday, June 11, 2008 10:22 PM
To: Robert W. Gensemer
Cc: Shephard.Burt@epamail.epa.gov; Humphrey.Chip@epamail.epa.gov;
PETERSON Jenn L
Subject: RE: Bioassay Interpretation

I think the question you pose about model performance is a good one and I do not have a good answer. I spent some time mapping out the 90% and 80% growth endpoints for Hyalella. Both of these thresholds were useful in delineating areas of contamination relative to know sources of contamination within Portland Harbor. However, according to John, Hyalella growth did not really correlate to any chemicals. It is curious as to why this would be the case if the data seems to map so well.

Regarding your first point, I assume that John was really just referring to data presentation. I expect that the risk characterization would evaluate the data according to the same thresholds. One thing that John mentioned in my conversation with him is they are looking at ways to get more out of the empirical data and seem to be relying more on the empirical data which I believe is also our first choice.

Eric

"Robert W. Gensemer" <rgensemer@parametrix.com>		To
06/11/2008 09:04 PM	Eric Blischke/R10/USEPA/US@EPA, Jennifer L Peterson <PETERSON.Jenn@deq.state.or.us>, Burt Shephard/R10/USEPA/US@EPA	cc
	Chip Humphrey/R10/USEPA/US@EPA	Subject
	RE: Bioassay Interpretation	

Eric: As a purely practical matter, this may be a reasonable compromise that allows us to break the stalemate and move forward. The empirical data analysis John proposes fits what we had in the problem formulation and as reiterated in further recent dicsussion. Not sure why they want to limit this to "color coding" instead of a formal risk characterization, but so long as the risk gradients are presented somewhere in a transparent fashion, I'm OK.

As for the models, John's arguments for "better fit" certainly appear more reasonable, but it does beg the question: If the empirical data can discriminate adverse effects less that 20-25%, why can't the models do so? Perhaps there are many good reasons for this I'm not tuned into. But I'm an empirically-biased person by default, so if that's the best the models can do and we're presenting the empirical data in full, then perhaps that's as far as we can go.

-Bob

-----Original Message-----
From: Blischke.Eric@epamail.epa.gov
[mailto:Blischke.Eric@epamail.epa.gov]

Sent: Wednesday, June 11, 2008 5:33 PM
To: Jennifer L Peterson; Shephard.Burt@epamail.epa.gov; Robert W. Gensemer
Cc: Humphrey.Chip@epamail.epa.gov
Subject: Bioassay Interpretation

I just got off the phone with John Toll. This is what we tentatively agreed to:

For the empirical data:

Anything statistically different than control will be flagged as a hit. Hits will be color coded as to magnitude of the hit (10%, 20%, 30%). This will be done for each of the four endpoints (CH10 survival, CH10 growth, HY28 survival, and HY28 growth)

For the predictive models:

We will apply the RSET approach to the model development. The following criteria will be used: CH10 and HY28 survival: SL1 = 10%, SL2 = 20%; CH10 growth: SL1=20%, SL2 = 30%; HY28 growth: SL1 = 25%, SL2 = 40%. As with RSET, if any one of the four SL2 values are exceeded or any 2 of the SL1 values are exceeded, the data will be presented as a hit. Empirical data will be presented against these criteria as well.

The rationale for this approach is:

1) The empirical data collected at the PH site is of high quality and is our strongest line of evidence. This approach ensures that we will get the most out of the empirical data set.
2) It is difficult to tease out hits with the predictive models. This approach provides a higher threshold for documenting a hit in the predictive models and is consistent with the RSET approach.

John is running this by the LWG team. I am running this by you three. Please let me know your view on this approach at your earliest convenience.

Thanks, Eric